International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 5672-5684

https://doi.org/10.70517/ijhsa464407

Application of Deep Convolutional Networks in the Identification of Microscopic Morphology of Traditional Chinese Medicine Herbs

Hanying Wang[1,2], Zhi Chen[1,2], Jiabo Huo[1] and Xingguo Han[1,2,*]
[1] Guangxi Key Laboratory of Special Engineering Equipment and Control, Guilin University of Aerospace Technology, Guilin, Guangxi, 541004, China
[2] School of Mechanical Engineering, Guilin University of Aerospace Technology, Guilin, Guangxi, 541004, China

Corresponding authors: (e-mail: Hxingg2010@163.com).

Abstract With the continuous development and expansion of the Chinese herbal medicine market, the quality control of herbs has become particularly important. In this study, a micro-morphological feature recognition method based on deep convolutional neural network (CNN) for the traditional Chinese herbal medicine maitake is proposed. First, images of different grades of maitake were collected, and multiple features such as morphology, color, and texture were extracted after image preprocessing, feature extraction, and enhancement. Then, the extracted images were classified and recognized using a deep learning algorithm. The experimental results show that the top-1 accuracy of this paper's method on the test set is 97.14%, and the top-1 accuracy after migration learning is improved to 99.35%, and the macro accuracy reaches 99.54%. Compared with traditional algorithms combining image processing and machine learning, the method in this paper has significant advantages. In addition, the depth-separable convolutional structure effectively reduces the computational burden of the model. The method shows good application prospects in the quality identification of Chinese herbal medicine Maitong, and can provide powerful technical support for the quality control of Chinese herbal medicine.

Index Terms Deep convolutional neural network, Chinese herbal medicine, micro-morphological features, image processing, transfer learning, accuracy

I. Introduction

Chinese medicinal herbs are unique medicinal resources in China, and more than 14,000 species have been explored so far, which play a pivotal role in disease prevention and protection of people's health, and are precious national treasures of China [1]. Among them, rhizomatous medicinal plants account for more than 80% of all medicinal plant species, and have attracted much attention because of their important proportion and wide range of medicinal values [2]. However, with the expansion of China's international market, traditional Chinese medicine resources are frequently patented and occupied by other countries. Taking Japan as an example, Japanese herbal formulas based on classical Chinese medicinal works such as “The Treatise on Typhoid Fever” and “The Essentials of the Golden Chamber” have surpassed China to become the authoritative representative of international recognition in just over 30 years since the 1970s because of its higher attention to scientific research, raw drug sources, quality control and patent protection than that of China [3]-[5]. Therefore, to improve international competitiveness and national soft power, it is imperative to strengthen the scientific research of Chinese herbal medicines, refine the production process, and improve the quality of Chinese herbal medicines [6].

Roots and stems of Chinese herbal medicines have different shapes and complex compositions [7]. Fresh herbs have high internal water content, and due to their thick structure, they usually require a long time for drying and processing, even up to two weeks for some herbs [8]-[11]. In addition, improper removal of water during the drying process can easily lead to the corruption, deterioration and discoloration of herbs, and even the phenomenon of insects and ants [12]. Excessively high temperatures can also cause an imbalance in the internal moisture gradient of Chinese herbal medicines, leading to browning reactions, resulting in the loss of active ingredients, and affecting the overall quality level of the Chinese herbal medicines and their clinical use [13]. Therefore, it is imperative to characterize the micro-morphology of traditional Chinese herbal medicines.

In this paper, a micro-morphological feature recognition method based on deep separable convolutional neural network (DSC) for Chinese herbal medicine is proposed. The method combines traditional image processing techniques and modern deep learning algorithms, and realizes automated recognition and classification of images of Chinese herbal medicines through preprocessing, feature extraction, and model training of image data. In the experiment, for Maitong, a common Chinese herbal medicine, this paper extracts morphological features, color

features and texture features from various aspects respectively, and classifies them by combining with the deep learning model. The experimental results show that the method in this paper is better than the traditional algorithm in terms of recognition accuracy and computational efficiency.

The core innovation of this study is the introduction of DSC, a structure that has lower computational complexity compared to traditional CNNs while ensuring recognition accuracy. In the future, with the continuous improvement of the dataset and the optimization of the model, the deep learning-based recognition technology of Chinese herbal medicine will have a wide range of application prospects, especially in the automated classification and quality detection of Chinese herbal medicine.

II. Material and morphological characterization methods

II. A. Selection and image acquisition of traditional Chinese medicine herbs

The main purpose of this study is to identify the superior, intermediate and inferior grades of graded traditional Chinese medicinal herbs, and Maitake Maitake was used as the study material. The black flannel was used as the acquisition background, a number of different grades of maitong were randomly selected, and a mixture of maitong was randomly flattened onto the black flannel, and a camera with 12 megapixels was utilized to vertically capture the images of maitong under sufficient light without shadows, and the images of the seeds with good acquisition effect were selected for image preprocessing.

II. B. Image Preprocessing

II. B. 1) Image Grayscaling Processing

Currently commonly used grayscale processing method is to take the weighted average method, this method is new R, G B three-component value will take the average of the different weights. The formula used is shown in (1):

$$Gray = 0.3R + 0.59G + 0.11B \tag{1}$$

In the above equation: Gray is used to represent the gray scale. R, G, B denote the red, green, and blue channels, respectively, and 0.3, 0.59, and 0.11 are the weighting coefficients of the three channels, respectively.

II. B. 2) Image Filtering

In this study, we choose the median filter processing with fast operation speed, which is a very classical method of smoothing noise reduction, removing the effect of noise while also protecting the edge information and image details of maitake. The median filtering algorithm sorts the number of n in a small-to-large manner, as shown in Equation (2):

$$y = \begin{cases} x_{i\left(\frac{n+1}{2}\right)} & n \in \{1,3,\cdots 2k+1\} \\ \dfrac{x_{i\left(\frac{n+1}{2}\right)} + x_{i\left(\frac{n}{2}\right)}}{2} & n \in \{2,4,\cdots 2k+1\} \end{cases} \tag{2}$$

In the above equation, y is the median value of different x, i is the center position of the window, and k is a positive integer. The median filtering process is to replace the value of a point in the McDonald's image with the median value of the points in the neighboring domain, so that the value of the surrounding pixels is close to the true value, so as to achieve the purpose of eliminating the noise, as shown in Equation (3):

$$F(x_0, y_0) = \left[\underset{(x,y)\in S}{sort} F(x,y) \right]_{\frac{N+1}{2}}, N \geq 0 \tag{3}$$

In the above equation (3), $F(x_0, y_0)$ denotes the median value of the image gray scale, S is the set of neighboring domains of the pixel (x_0, y_0), N denotes the number of elements in the set S, and sort denotes the sorting mode of the function.

II. B. 3) Image enhancement

After the image has been filtered and noise reduced, the edges of the binarized image of the target become somewhat smooth, resulting in some detailed features that are not very obvious, so it is also necessary to enhance the image. Enhancement is performed using histogram equalization, and the algorithms for enhancement have two different forms based on the frequency domain and based on the spatial domain. Among them, frequency domain based is an indirect enhancement by correcting the values of the image transform coefficients in their corresponding

transform domains [14]. While based on the spatial domain is the direct grayscale operation on the image, which is subdivided into point operation processing and neighborhood enhancement processing.

II. B. 4) Image segmentation

Threshold segmentation is a classical segmentation algorithm, mainly from the gray scale to compare the difference between the target and background, so that the pixel level is divided into a number of classes by different thresholds, so as to achieve the effect of separating the target and the background. OTSU is a method of automatically selecting the threshold value according to the gray scale characteristics of the image by maximizing the interclass variance, so as to make the target and the background divided into two parts.

II. B. 5) Image Morphology Processing

The formula definitions of both open and closed operations are two very useful quadratic operations in morphology, and the definitions of both are shown in Eqs. (4) and (5), respectively:

$$A \circ B = (A\,!\;\; B) \oplus B \tag{4}$$

$$A \cdot B = (A \oplus B)\,!\;\; B \tag{5}$$

The above equations A and B denote the image and structural elements to be processed, respectively. $!$ denotes corrosion. $\oplus$ denotes expansion.

II. C.Deep Learning Sample Expansion

The individual seed images extracted above were expanded with samples. The expansion was done in such a way that the biological characteristics remained as unchanged as possible.

II. D.Extraction of morphological features

II. D. 1) Extraction of the perimeter of the maitake

Perimeter is used to rate the size of the object's appearance and shape, which can be expressed as a specific numerical value of the length of the object's contour. Usually, the 8-neighborhood connectivity rule is applied to extract the boundary points of the object sequentially using a specific direction, and the number of pixels with 4- and 8-connectivity between the extracted point and the next extracted point is recorded during the extraction, 4-connectivity means that the pixels are connected horizontally or vertically, and its number is denoted by N_a. 8-connectivity means that the pixels are connected in a diagonal manner, and its number is denoted by N_d. It is used as shown in equation (6).

$$P = N_a + \sqrt{2} N_d \tag{6}$$

II. D. 2) Extraction of Maitake area

Area is a unit that indicates how much of a plane a target occupies, and is a measure that rates the size of a target in two dimensions. The area of the McDonald's can be obtained by counting the total number of pixel points in the area. Assuming that each pixel represents one unit of area imaged, the area can be found by counting the sum of the number of pixel points in the occupied area. The area is represented by the symbol “S” as shown in equation (7).

$$S = \sum_{(x,y)\in\Omega} 1 \tag{7}$$

II. D. 3) Extraction of roundness of maitake

Roundness indicates the degree to which an object's contour deviates from the circular standard and reflects, within certain limits, the tightness of the area, as determined by area and circumference. Roundness is used as a rating indicator. When the shape of the object contour is round, the value of roundness is the maximum value of 1, and the degree of deviation from the circle in other shapes, the value of roundness is less than 1. The degree of roundness is expressed by the symbol “B”, as shown in formula (8).

$$B = \frac{4\pi A}{P^2} \tag{8}$$

II. D. 4) Extraction of rectangularity of maitake

Rectangularity is the smallest rectangle external to the object and indicates how full the object is in the rectangle. The placement angle of the target object has a large effect on the external rectangle. The degree of rectangularity is represented by the symbol “R”, as shown in equation (9).

$$R = \frac{S}{d_{\min} \times d_{\max}} \tag{9}$$

In Eq. (9) $d_{\min}$, $d_{\max}$ denote the short axis, and the long axis, respectively.

II. D. 5) Extraction of Maitake elongation length

The elongation length describes the slenderness of the maitake, that is, the degree of elongation of the maitake form. It is represented by the symbol “E”, as shown in equation (10).

$$E = \frac{\min(h, w)}{\max(h, w)} \tag{10}$$

In equation (10) h and w denote the height, width respectively.

II. E. Extraction of color features

The RGB color model is a color model structure based on the Cartesian coordinate system. In the coordinate system, red (R), green (G), and blue (B) represent the three axes, and any one color can be corresponded to the sum of the three base colors of a particular component, and the RGB color model is shown in Figure 1. In the RGB color space, the three primary colors can represent color and luminance.

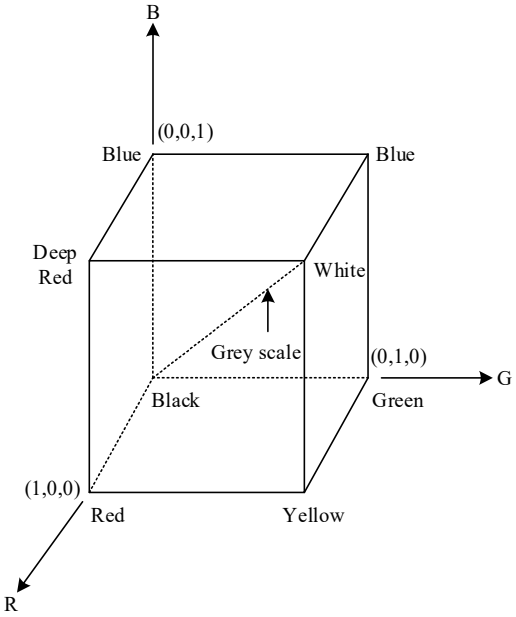

Figure 1: the model of RGB color

Compared with RGB model HSV model is a color model with strong intuition. This model classifies the collected color information according to the three attributes of hue (H), saturation (S), and lightness (V), and the conversion equations between RGB and HSV are shown in (11) to (14).

$$H = \cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\} \tag{11}$$

$$H = 360^{\circ} - \cos^{-1}\left\{\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right\} \tag{12}$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \tag{13}$$

$$V = \frac{1}{3}(R+G+B) \tag{14}$$

II. F. Extraction of texture features

Texture is a visual feature that can bring out the same characteristics of an image and describes the surface nature of an object image. In this study, the texture features of Maitake are extracted from four feature parameters: smoothness, moment of inertia, consistency, and entropy. The calculated formulas are shown in (15) to (18).

$$R = 1 - \frac{1}{(1+\sigma^2)} \tag{15}$$

$$\mu_2 = \sum_{I=0}^{L-1}(Z_I - M)^2 P(Z_I) \tag{16}$$

$$U = \sum_{I=0}^{L-1} p^2(Z_i) \tag{17}$$

$$e = -\sum_{i=0}^{L-1} p(Z_i)\log_2 p(Z_i) \tag{18}$$

where, Z_I denotes the random value of luminance. $P(Z_I)$ denotes the histogram of gray levels. L denotes the number of levels of gray scale.

III. Recognition of micro-morphological features of traditional Chinese herbal medicines based on deep learning

III. A. Block diagram of the overall algorithm

The micro-trait identification model of traditional Chinese herbal medicine Maitake is shown in Fig. 2. The algorithm is mainly divided into four parts: image fusion, ZCAization, deep separable convolutional neural network model training and classification. Firstly, image fusion, after data augmentation of the acquired traditional Chinese herbal medicine images, image fusion based on homomorphic filtering and CNN is carried out to obtain the microtrait images of Chinese herbal medicines. Secondly, ZCAization, the processed images are sequentially processed so that the correlation between the data is reduced. Then there is a deep separable convolutional neural network to train the processed herbal microtrait images to obtain data features. Finally, classification is performed to identify the herbs by Softmax classifier.

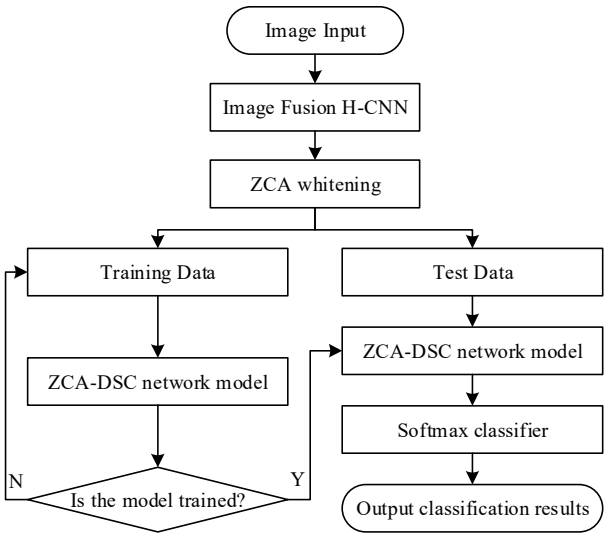

Figure 2: Identification model of Chinese medicinal materials microtraits

III. B. Design of Algorithm for Identification of Traditional Chinese Medicine Maitake

III. B. 1) Image Fusion

Firstly, homomorphic filtering is used to deal with the image quality problems caused by light intensity, brightness and image contrast in the image acquisition process of Chinese herbal medicines. Secondly, the processed images are fused based on CNN to obtain the microtrait images of Chinese herbal medicines. The image model of homomorphic filtering is constructed according to the reflection theorem, and the model illumination function $f(x,y)$ is obtained by multiplying the irradiation component $i(x,y)$ with the reflection component $r(x,y)$ as shown in equation (19). Among them, the change process of the irradiation component is relatively slow, which corresponds to the low-frequency part. The reflection component of the change process is the opposite, equivalent to the high-frequency part of its nature by the imaging object surface characteristics of the decision.

$$f(x,y) = i(x,y) \cdot r(x,y) \tag{19}$$

To make the computation of the model more convenient, a logarithmic transformation is applied to Eq. (19), which transforms to an additive computation in the logarithmic domain.

$$\ln(f(x,y)) = \ln(i(x,y)) + \ln(r(x,y)) \tag{20}$$

The model is then transformed into the frequency domain by Fourier transform, i.e:

$$\ln(f(x,y)) = \ln(i(x,y)) + \ln(r(x,y)) \tag{21}$$

The $F(\ln f(x,y))$ obtained after logarithmic and Fourier transformations of the light function $f(x,y)$ by homomorphic filtering $H(u,v)$ is processed to obtain $S(u,v)$ as shown in equation (22):

$$S(u,v) = H(u,v) * F(\ln f(x,y)) \tag{22}$$

Second, the obtained $S(u,v)$ is then inverted to obtain $s(x,y)$:

$$s(x,y) = F^{-1}(S(u,v)) \tag{23}$$

Finally, exponentiate $s(x,y)$ to obtain the image processing result $f'(x,y)$ after this model.

$$f'(x,y) = \exp(s(x,y)) \tag{24}$$

In order to obtain microtrait images of Chinese herbal medicines, in this chapter, the homomorphic filtered processed images are fused using a CNN-based deep fusion algorithm to convert the multifocused image fusion into a binary classification problem.

The CNN model used in the H-CNN-based fusion algorithm is shown in Fig. 3, and its feature extraction as well as classification processes replace the activity level measurement and fusion rule design processes in the image fusion process, respectively.

1) Focusing level measure

The images H_A and H_B are firstly obtained by homomorphic filtering and grayscaling operations, and then H_A and H_B are inputted into the trained CNN network in order to obtain the focus degree metric map S. Where the values in S are all in the range of [0,1], indicating that the focusing characteristics of the input images H_A and H_B are complementary, and the closer their values converge to 0 or 1, the more the complementary parts of the input images H_A and H_B are focused. In order to obtain a focus map M with the same size as the input image, the values in S are assigned to all pixels in the complementary part of the focus map M.

2) Binary Focus Map

In order to retain as much useful information as possible, the focus map M is further processed using the "select-max" strategy. The threshold value is set to 0.5 to obtain a binary focus map T based on the focus map M. This is shown in equation (25).

$$T(x,y) = \begin{cases} 1, & M(x,y) > 0.5 \\ 0, & otherwise \end{cases} \tag{25}$$

where x and y are the coordinate values of the pixel points.

3) Consistency verification

Since some misclassified pixels may appear in the process of acquiring the binary focus map T, a small region removal strategy is utilized to deal with the misclassified pixels in the binary focus map T, i.e., the region whose area is less than the threshold value of $0.01\times H\times W$ is inverted with respect to the values in its binary focus map T. Where H and W are the height and width of the source image, respectively. Meanwhile, bootstrap filtering is used to further optimize the binary focus map T to obtain the decision map $D(x,y)$.

4) Fusion

Finally, the fused image F is obtained by pixel weighted averaging method using the decision map $D(x,y)$.

$$F(x,y)=D(x,y)H_A(x,y)+(1-D(x,y))H_B(x,y) \tag{26}$$

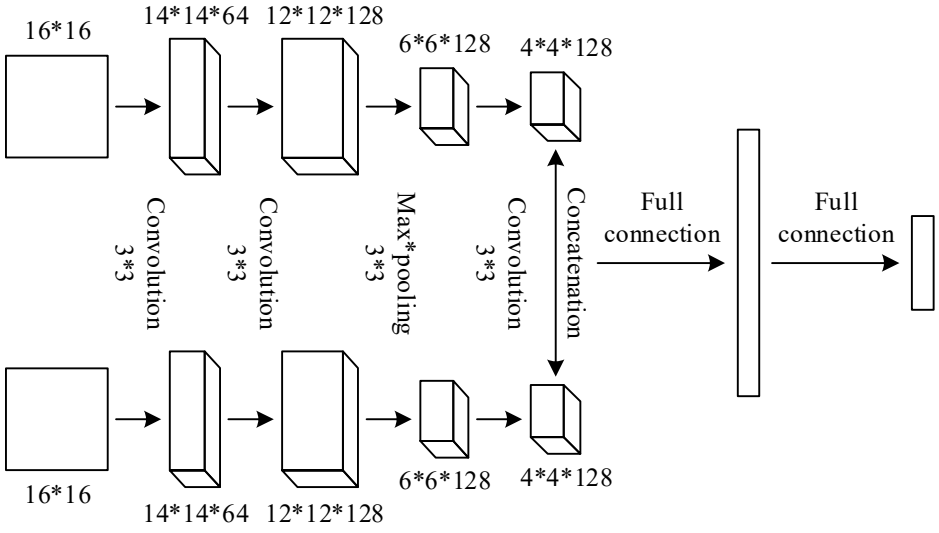

Figure 3: CNN model based on the integrated CNN fusion algorithm

III. B. 2) Data enrichment

In deep learning, the network is complex and has more parameters, which leads to a larger amount of data required in its training process, and in order to make the obtained model more realistic as well as to avoid overfitting, it is therefore necessary to augment the acquired microtrait images of the herbs with data.

III. B. 3) ZCA whitening

As a common method to reduce data correlation and redundancy, ZCA whitening is implemented by transforming the data matrix, keeping the variance of the features and the covariance of the data unchanged. In this section, ZCA whitening is selected to process the microtrait images of Chinese herbal medicines to reduce their data redundancy. Take the Chinese herbal medicine microtrait image dataset $\{x^{(1)},x^{(2)},\ldots x^{(m)}\}$ with sample dimension n as an example for illustration, the steps are as follows:

1) First, the input herbal microtrait image data $\tilde{x}^{(i)}$ is normalized in terms of brightness and contrast. That is, for all pixel points $x^{(i)}$, the size of the ratio of its difference from the mean value $mean(\tilde{x}^{(i)})$ and its standard deviation.

Since a zero denominator is meaningless, a constant $e=10$ is added to the denominator, i.e.:

$$x^{(i)}=\frac{\tilde{x}^{(i)}-mean(\tilde{x}^{(i)})}{\sqrt{\mathrm{var}(\tilde{x}^{(i)})+e}} \tag{27}$$

2) Then, the covariance matrix Σ is computed for the training samples, and since the data are not completely uncorrelated with each other, the resulting covariance matrix Σ is not a diagonal array, as shown in (28).

$$\Sigma=\frac{1}{m}\sum_{i=1}^{n}(x^{(i)})(x^{(i)})^T \tag{28}$$

3) Second, transform Σ to obtain its diagonal array D as well as the eigenvector V, which is:

$$[V,D]=eig(\Sigma) \tag{29}$$

4) Finally, ZCA whitening is performed and the calculation formula is shown in (30).

$$\hat{x} = V(D + \varepsilon_{ZCA} I)^{-1/2} V^T \tag{30}$$

where ε_{ZCA} is a constant with a value of 0.01 and I is the unit matrix.

The ZCA whitening operation is used to map the input vector x into the x_{ZCA} space so that its variance is consistent in all dimensions, and the data correlation and redundancy of the microtrait images of Chinese herbal medicines are reduced to some extent.

III. B. 4) The ZCA-DSC model

This section classifies and identifies microtrait images of Chinese herbal medicines based on depth-separable convolution, which is less time-consuming and relatively simple in structure than traditional convolutional neural networks. The depth-separable convolutional layer is decomposed on the basis of the traditional standard convolutional layer, which consists of a depth-by-depth convolution as the input channel of image features and a 1□1 standard convolution that fuses the depth-by-depth convolutional features, which can make the model structure simpler to a certain extent.

The input feature map F of the standard convolutional layer with the size $D_F \times D_F \times M$ is processed by the standard convolution operation K with the size $D_K \times D_K \times M \times N$, and the output feature map Y with the size $D_Y \times D_Y \times N$ is obtained. The corresponding computational quantities and parameter quantities are shown in Eq. (31) and Eq. (32), respectively.

$$D_K \times D_K \times M \times N \times D_F \times D_F \tag{31}$$

$$D_K \times D_K \times M \times N \tag{32}$$

where M and N are the number of input and output feature channels, respectively, and $D_K \times D_K$ is the size of the convolution kernel.

The conventional standard convolutional layer is split into a depth-by-depth convolution of size $D_K \times D_K \times M$ and a point-by-point convolution of size $D_K \times D_K \times N$ of 1×1, which form the depth-separable convolutional layer. Among them, the depth-by-depth convolution is mainly a filtering operation on the features, which is computationally intensive:

$$D_K \times D_K \times M \times D_F \times D_F \tag{33}$$

1×1 point-by-point convolution is responsible for the conversion of the channel, by linearly combining with the depth-by-depth convolution, new features are obtained, so that the depth-separable convolution layer completes its entire process, then the computational and parametric quantities of the depth-separable convolution layer are shown in equation (34) and equation (35), respectively.

$$D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F \tag{34}$$

$$D_K \times D_K \times M + M \times N \tag{35}$$

From Eq. (36) and Eq. (37), it can be seen that the computational and parametric quantities of the depth-separable convolutional layer are reduced in the same proportion compared to the conventional standard convolutional layer.

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2} \tag{36}$$

$$\frac{D_K \times D_K \times M + M \times N}{D_K \times D_K \times M \times N} = \frac{1}{N} + \frac{1}{D_K^2} \tag{37}$$

IV. Experimental results and analysis

IV. A. Experiments and analysis of traditional Chinese herbal medicine identification

IV. A. 1) Model performance validation

In this set of experiments, ResNet-50, DenseNet-169, MoblieNetV2, efficientNet-B0, and the model in this paper were used to train the TCM image dataset without and with migration learning. The L3-L9 modules of Attention-

TCM-Net are loaded with the convolutional kernel parameters of EfficientNet-B0 trained on ImageNet dataset as initial parameters, and the other layers of the network are still randomly initialized. The overall parameters of the network are used in the training process to better adapt to the image features and attention mechanisms of TCM. The comparison of the test set results is shown in Table 1, and the test results are synthesized and evaluated using three-fold cross-validation. As can be seen from the table, the model based on this paper's algorithm is significantly higher than the algorithm combining traditional image processing and machine learning in traditional Chinese medicine classification accuracy, indicating that deep learning algorithms are more advantageous in Chinese medicine identification. At the same time, migration learning can effectively improve the performance of the network, and the top-1 accuracy and macro accuracy of the network structure are both improved, in which the top-1 accuracy is improved by more than 2%. Therefore, it is necessary to load the pre-trained convolutional kernel parameters, which will help to improve the accuracy and convergence speed of the network. Comparing the experimental results of all the networks, the method in this paper achieves the best results on the traditional Chinese medicine image test set.Compared with the rest of the best Efficient-B0 model results, the mean value of the top-1 accuracy of the tri-fold cross-validation of this paper's method reaches 97.14% with random parameter initialization, and the mean value of the top-1 accuracy in the case of migration learning reaches 99.35%. In terms of macro-accuracy, the method of this paper even reaches 99.54%, achieving better results for each class of traditional Chinese medicine, indicating that the method of this paper has good recognition ability for the image dataset of traditional Chinese medicine constructed in this paper, which is more advantageous than other networks. In addition, compared with the classic CNN network models such as ResNet and DenseNet, the method in this paper uses deep separable convolution, the model space complexity and time complexity is greatly reduced, the number of parameters is 3.98M, and the FLOPs are 0.33B, and the introduction of the CBAM and the related attention module does not increase the computational burden of the model, so that we can guarantee the good performance of the network and still maintain the lightweight. The lightweight design is still maintained under the premise of ensuring good performance of the network.

Table 1: The Comparison of test set results

	Nonmigration learning		Migration learning		Image size	Parameter quantity	FLOPs
	The average rate of accuracy	Macro accuracy mean	The average rate of accuracy	Macro accuracy mean			
ResNet-50	96.06	96.82	98.45	99.19	240×204	23.51M	4.14B
DenseNet-169	96.39	97.11	98.63	98.88	240×204	12.41M	3.49B
MobileNetV2	95.87	96.64	97.39	98.09	240×204	2.15M	0.3B
Efficient-b0	95.98	97.29	98.45	99.26	240×204	4.05M	0.43B
Ours	97.14	97.62	99.35	99.54	240×204	3.98M	0.33B

IV. A. 2) Analysis of the identification of specific categories

In this group of experiments, ResNet-50, DenseNet-169, MoblieNetV2, and efficientNet-B0 were used to test each type of traditional Chinese medicines in the test set in turn using the method of this paper, and then the comparison between learning without transfer and learning with transfer was carried out. Finally, the confusion matrix and ROC curves are plotted to explore the difficulty of recognizing each class of Chinese medicines, and the ease of confusion. The comparison of the accuracy rate of each traditional Chinese medicine without transfer learning is shown in Table 2, and the comparison of the accuracy rate of each traditional Chinese medicine with transfer learning is shown in Table 3. The results show that in both cases, several networks are able to recognize hawthorn and half-summer better because the images of hawthorn and half-summer are quite different from the color, such as hawthorn charcoal and raw hawthorn, half-summer and ginger half-summer. The average recognition accuracy of hawthorn was also kept above 99% at one time, with burnt hawthorn and hawthorn charcoal being the easiest to distinguish. On the other hand, Chuanbeimu is prone to confusion when pre-training parameters are not used due to more similar visual features such as color and shape. Among them, Furnace Shellfish has the worst recognition ability, and the highest of all models is the method of this paper, which has an accuracy of 88.8%, while Green Shellfish has an accuracy of 93%, and Pine Shellfish has an accuracy of 98.2%, which indicates that among the three types of Shellfish, Furnace Shellfish is the most difficult to differentiate, while Pine Shellfish has more obvious features and a higher relevant recognition rate.

After the improvement of migration learning, the convolutional kernel feature extraction ability of the five networks was enhanced, and the recognition accuracy of Sichuan shellfish achieved a significant rise, and the method in this paper obtained the best recognition results, in which the accuracy of furnace shellfish rose to 98.1%, green shellfish to 98.2%, and pine shellfish to 98.6%, and the networks performed well.

Table 2: Comparison of accuracy rates without transfer learning

Network model	Fried hawthorn	Carbonized Crataegus	Crataegus Crataegus	Fresh hawthorn	Furnace bay	Green shellfish	Pine bay	semisummer	Middle summer	Pinellia Ternata	Ginger half summer
Resnet50	99.7	99.7	99.8	99.8	87.4	89.8	96.6	96.8	99.2	98.2	99.1
Densenet169	99.6	99.7	99.7	99.9	88.4	90.7	95.9	97.2	99.3	98.1	99.4
MoblieNetV2	99.5	100	99.8	99.6	86.8	89.1	96.9	96.6	98.4	98.4	99.2
Efficient-b0	99.5	99.6	100.1	99.6	88.4	90.9	96.6	97.2	99.3	99	100.1
Ours	99.6	99.3	100	99	88.8	93	98.2	98.7	99.9	99	99.6

Table 3: Comparison of traditional with migration learning

Network model	Fried hawthorn	Carbonized Crataegus	Crataegus Crataegus	Fresh hawthorn	Furnace bay	Green shellfish	Pine bay	semisummer	Middle summer	Pinellia Ternata	Ginger half summer
Resnet50	99.7	100	100	100	96.1	95.9	97.7	99.3	100	99.5	100
Densenet169	99.8	100	100	99.9	96.7	94.8	98.4	99.8	100	100	100
MoblieNetV2	99.8	100	100	100	94.2	95.5	97.3	99.3	100	99.1	100
Efficient-b0	99.9	100	100	100	97.8	94.2	98.5	100	100	100	100
Ours	99.9	100	100	100	98.1	98.2	98.6	99.9	100	100	100

IV. B. Recognition analysis of micromorphological features of Maitake images

IV. B. 1) Feature extraction of images

(1) Geometric shape and color characteristics of herbs

The images of herbs were converted into binary images, the area, perimeter, and length-short axis of the target were extracted, and the shape parameters such as rectangularity, circularity, and length-short axis ratio were calculated. The data were examined by SPSS using Kruskal-Wallis. The measurement results of the shape and color parameters of the herbs are shown in Table 4 (different superscript letters of the data in the same column in the table indicate that the differences are statistically significant ($P < 0.05$)).The H (hue) and S (saturation) values of the HSV color space are not much affected by the light intensity, and can be used as a parameter for describing the color of the object. The images of herbs were transferred to HSV space and the mean values of H and S were extracted. The data were examined by Kruskal-Wallis method.

Table 4: The shape and color parameters of the medicinal materials are measured

Medicinal materials	Rectangle degree	Roundness	Length axis ratio	Tonal mean	Saturation mean
Chenmatin	0.74±0.02a	2.02±0.24a	3.66±0.66a	35.21±1.66a	0.53±0.023a
Cumin	0.72±0.05a	1.92±0.29a	3.53±0.54a	38.85±1.52b	0.51±0.031b
Santas	0.54±0.16b	3.06±0.74b	5.53±1.21b	40.9±1.63c	0.47±0.05c

(2) Texture features of herbs

The grayscale co-production matrix method is a common method to extract the texture features of the image. The image of herbs was converted to gray scale image, and the gray symbiotic matrix of the image and the

corresponding parameters such as contrast, correlation and homogeneity were calculated by graycomatrix function and graycoprops number respectively (selecting the number of gray levels to be 64, the pixel offset distance to be 1, as well as the four directions of 0°, 45°, 90°, and 135°), and then the values of each parameter in the four directions were summed up as the measured value of that parameter. The images of the herbs were also converted to HSV color space, and the corresponding texture parameters for the H and S components were calculated under the same conditions, respectively. The data were examined by Kruskal-Wallis method. The measured results of the texture parameters of the herbs are shown in Table 5.

Table 5: The texture parameter measurements of medicinal materials

Medicinal materials	Grayscale texture			Tonal texture			Saturation texture		
	Contrast ratio	correlation	homogeneity	Contrast ratio	correlation	homogeneity	Contrast ratio	correlation	homogeneity
Chenmatin	49.66±10.83a	3.91±0.01a	2.69±0.26a	70.17±14.63a	4±0.01a	3.51±0.02a	25.28±7.86a	3.84±0.02a	2.94±0.26a
Cumin	41.99±7.15b	3.92±0.01a	3.29±0.12a	75.19±13.02b	3.9±0.02a	3.5±0.09a	19.36±3.42b	3.98±0.01b	3.19±0.13b
Santas	30.78±8.59c	3.9±0.01b	3.51±0.16c	63.38±16.82b	4.06±0.02b	3.64±0.03b	12.71±3.7c	4.24±0.01b	3.77±0.11c

(3) Area fraction of each tissue in cross section of herbs

The mid-section cross-section image of the sample mainly included epidermis, cortex, stone cell ring band, xylem and pith. According to the differences of each tissue in terms of gray value, wall thickness, size, etc., the sample image was denoised, and then different thresholds were applied to gradually segment the image into a number of sub-regions. The regionprops function was used to calculate the area of each segmented part and its filled part to obtain the epidermal area S1, the area of the stone cell annulus S2, the area of the xylem S3, the area of the pith S4, the area of the mesocolumnus S5 (xylem + phloem), and the area of the transverse section S6. and calculate the epidermal area fraction Ss1 = S1/S6, the area fraction of the area within the stone cell annulus (stone cell annulus + mesocolumn + pith) Ss2 = (S2 + S5 + S4)/S6, the medullary area fraction Ss3 = S4/S6, the area of xylem plus medullary area over the area within the stone cell annulus Ss4 = (S3+S4)/(S2+S5+S4), and the area of pith over the xylem plus medullary area Ss5=S4/(S3+S4). The data were tested by Kruskal-Wallis method. The area fraction measurements of each tissue of the herbs are shown in Table 6.

Table 6: The results of the measurement of the surface integral of the herbs

Medicinal materials	Ss1	Ss2	Ss3	Ss4	Ss5
Chenmatin	0.047±0.005a	0.02±0.003a	0.009±0.003a	0.67±0.06a	0.578±0.08a
Cumin	0.051±0.009a	0.012±0.002a	0.002±0.002a	0.668±0.02a	0.561±0.06a
Santas	0.047±0.009a	0.011±0.001b	0.0005±0.0005b	0.49±0.052b	0.102±0.06b

IV. B. 2) Principal Component Analysis and Pattern Recognition

(1) Principal component analysis

The raw data (n = 50) of the 19 eigenparameters in Tables 4 to 6 were collapsed into a data matrix (50 × 19) and subjected to principal component analysis using the princomp function to obtain the loadings matrix (19 × 19), the data matrix represented by the principal component space (50 × 19), and the eigenvalues of the covariance matrix or the principal component variance (19 × 1). The principal component variances (eigenvalues) were organized. The principal component variances and variance contributions are shown in Table 7. The top 10 components with a cumulative contribution rate of variance ⩾95% were selected as input data for pattern recognition.

Table 7: Principal component variance and variance contribution

Constituent	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Eigenvalue	10.32	2.12	1.3	1.23	0.89	0.73	0.6	0.41	0.31	0.28	0.16	0.13	0.12	0.11	0.12	0.11	0.04	0.02	0.01

Variance contribution	54.29	11.15	6.84	6.47	4.68	3.84	3.16	2.16	1.63	1.47	0.84	0.68	0.63	0.58	0.63	0.58	0.21	0.11	0.05
Cumulative contribution of variance	54.29	65.44	72.28	78.75	83.43	87.27	90.43	92.59	94.22	95.69	96.53	97.21	97.84	98.42	99.05	99.63	99.84	99.95	100

(2) Pattern Recognition

MatLab2014a neural network pattern recognition toolbox was invoked, with 70% of the samples as the training set, 15% of the samples as the test set, and 15% of the samples as the validation set, the number of input nodes was 10 (10 principal components), the number of nodes in the output layer was 3 (3 herbs), and the output vectors [100] stood for zygomatic maitake, [010] stood for Chuan maitake, and [01] stood for Shan maitake. The number of hidden layer nodes is greater than the sum of input nodes and output nodes, which is set to 15 here. The training set is used to build the model, and the validation set is used to describe the generalization performance of the model, and the training is stopped when the generalization performance of the model stops growing. The recognition results of this paper's method are shown in Table 8. From the table, the combined recognition rate of this paper's method is 94.2%, which has good recognition results.

Table 8: Recognition results

Categories	Training set			Test set			Verification set			Synthesize		
	Quantity	Identification number	Recognition rate (%)	Quantity	Identification number	Recognition rate (%)	Quantity	Identification number	Recognition rate (%)	Quantity	Identification number	Recognition rate (%)
Chenmatin	35	31	88.6	9	7	77.8	11	10	90.9	55	48	87.3
Cumin	30	28	93.3	12	12	100	6	6	100	48	46	95.8
Santas	38	38	100	8	8	100	7	7	100	53	53	100
Synthesize	103	97	94.2	29	27	93.1	24	23	95.8	156	147	94.2

V. Conclusion

The model proposed in this paper has achieved remarkable results in traditional Chinese herbal medicine identification experiments. It is verified through experiments that the micro-morphological feature recognition method of Chinese herbal medicines based on deep separable convolutional neural network has strong application value. Under the condition of no migration learning, the Top-1 accuracy is 97.14% and the macro average accuracy is 97.62%; under the condition of migration learning, the Top-1 accuracy is increased to 99.35% and the macro average accuracy is 99.54%. Compared with other classical CNN models, the amount of model parameters of this paper's method is reduced to 3.98M, the computational complexity is reduced, and the FLOPs are 0.33B, demonstrating higher computational efficiency and lower resource consumption. In addition, for different kinds of Chinese herbs, the method in this paper can also effectively distinguish them, with strong generalization ability. By further optimizing the model and expanding the dataset, this method is expected to be more widely applied in the field of quality control and identification of Chinese herbal medicines in the future.

Funding

This research was funded by the institution of Middle-aged and Young Teachers' Basic Ability Promotion Project of Guangxi (Grant No.2021KY0802); Guangxi Nature Science Foundations (Grant No.2024GXNSFAA010338); and GUAT Special Research on the Strategic Development of Distinctive Interdisciplinary Fields (Grant No.TS2024121).

References

[1] Wang, W. Y., Zhou, H., Wang, Y. F., Sang, B. S., & Liu, L. (2021). Current policies and measures on the development of traditional Chinese medicine in China. Pharmacological research, 163, 105187.

[2] Odey, M. O., Iwara, I. A., Udiba, U. U., Johnson, J. T., Inekwe, U. V., Asenye, M. E., & Victor, O. (2012). Preparation of plant extracts from indigenous medicinal plants. International Journal of Science and Technology, 1(12), 688-692.
[3] Kuchta, K. (2019). Traditional Japanese kampo medicine–history of ideas and practice; Part 1: from ancient shamanic practice to the medical academies of Edo. Traditional & Kampo Medicine, 6(2), 49-56.
[4] Liu, S., Matsuo, T., Matsuo, C., & Abe, T. (2022). Traditional Chinese medicines and prescriptions brought from China to Japan by a monk (Jianzhen, Japanese: Ganjin): A historical review. Compounds, 2(4), 267-284.
[5] Zhu, G., Yan, H., Chen, L., Ren, Y., & Chu, G. (2019). Historical evolution of traditional medicine in Japan. Chinese Medicine and Culture, 2(1), 36-43.
[6] Xie, P. S., & Leung, A. Y. (2009). Understanding the traditional aspect of Chinese medicine in order to achieve meaningful quality control of Chinese materia medica. Journal of Chromatography A, 1216(11), 1933-1940.
[7] Brand, E., Leon, C., Nesbitt, M., Guo, P., Huang, R., Chen, H., ... & Zhao, Z. (2017). Economic botany collections: a source of material evidence for exploring historical changes in Chinese medicinal materials. Journal of Ethnopharmacology, 200, 209-227.
[8] Li, X., Wang, Y., Ma, X., Li, M., Kong, D., Phavady, P., ... & Abdelkader, T. K. (2024). Chinese medicinal materials' drying technologies advancements–Principles, energy performance, and influence on the bioactive components. Drying Technology, 42(12), 1815-1845.
[9] Gao, Y. (2025). Identification of common Chinese medicinal materials based on micro-morphological characteristics in traditional Chinese medicine pharmacies. Molecular & Cellular Biomechanics, 22(4), 1048-1048.
[10] Guo, X. H., Zhu, Y. Y., Wang, Z. Z., Guan, Y., Ding, X. F., Wang, W. J., ... & CHENG, Y. J. (2010). Comprehensive Characterization of Main Components and Microstructures in Five Different Species of Salvia Miltiorrhiza Bge. Spectroscopy and Spectral Analysis, 30(8), 2299-2302.
[11] Xu, Y., Song, W., Zhou, P., Li, P., & Li, H. (2015). Morphological and microscopic characterization of five commonly-used testacean traditional Chinese medicines. Acta Pharmaceutica Sinica B, 5(4), 358-366.
[12] Xue, Y., Liang, Z., Tan, H., Ni, L., Zhao, Z., Xiao, T., & Xu, H. (2016). Microscopic identification of Chinese medicinal materials based on X-ray phase contrast imaging: from qualitative to quantitative. Journal of Instrumentation, 11(07), C07001.
[13] Cheng, X., Ji, H., Cheng, C., Sun, Y., Cheng, H., Wang, D., ... & Liu, X. (2022). Comprehensive determination of the processing level of rhizome of Polygonatum sibiricum by macroscopic, micromorphological, and microscopic characterizations. Microscopy research and technique, 85(7), 2669-2678.
[14] Tonghuan Ran,Guangfeng Shi,Zhuo Zhang,Yuhao Pan & Haiyang Zhu. (2024). Hyperspectral Image Classification Method Based on Morphological Features and Hybrid Convolutional Neural Networks. Applied Sciences,14(22),10577-10577.